LA-UR-21-27930

Title: Detecting Electrical Anomalies via Overlapping Measurements

Author(s): Sontowski, Sina
Lawrence, Nigel Rhea
Deka, Deepjyoti

Intended for: Report

Issued: 2021-08-09

Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Detecting Electrical Anomalies via Overlapping Measurements

Sina Sontowski[1], Nigel Lawrence[2], and Deepjyoti Deka[3]

[1]sina@lanl.gov
[2]nlawrence@lanl.gov
[3]deepjyoti@lanl.gov

August 6, 2021

Abstract

As cyber-attacks against critical infrastructure become more frequent, it is increasingly important to be able to rapidly identify and respond to these threats. Therefore, we are investigating using multiple independent systems with overlapping electrical measurements to more rapidly identify anomalies. While prior research has explored the benefits of fusing measurements, the possibility of overlapping measurements from an existing electrical system has not been investigated. To that end, we explore the potential benefits of combining overlapping measurements both to improve the speed/accuracy of anomaly detection and to provide additional validation of collected measurements.

1 Introduction

In recent years, cyber attacks have become more frequent and increased in complexity and sophistication. The power grid is vulnerable to these attacks due in large part to an increase in inter-connectivity caused by Supervisory Control and Data Acquisition (SCADA) system modernization used to monitor and control parts of the power grid. Newer digital devices with increased connectivity can provide advantages for power grid operation but can also increase risk from potential cyber threats [8]. A cyber attack on the power grid can de-energize power system components and aggregate operating conditions by causing overloading and instability [7]. The 2015 Ukraine cyber attack exemplifies the threat that cyber attacks can pose to the power grid.

Many consumers in Ukraine temporarily lost power due to a cyber attack exploiting Windows vulnerability CVE-2014-4114. A phishing email was opened that installed BlackEnergy malware on the system. This vulnerability was not a zero-day vulnerability and the operator should have been aware of the vulnerability in the system and taken action beforehand [8].

While anomaly detection would not have prevented this attack, it would have been able to alert the operator about unusual activity in the system potentially allowing them to respond more rapidly. Thus, anomaly detection allows for an early detection of cyber-intrusions [7]. Anomaly detection identifies potential issues in the system and can therefore enable operators to respond to malicious activity as it happens. Therefore, it can be seen as an early warning mechanism [7].

In this paper, we are describing our approach to improve the speed and accuracy of anomaly detection by using overlapping electrical measurements from two independent systems taken at different points throughout Los Alamos National Laboratory's power grid. In Section 2, we explore related works and their significance. In Section 3, background information is provided for a better understanding of our implementation described in Section 4. A conclusion is provided and possible future work approaches are mentioned in Section 5 and Section 6, respectively.

This paper's contributions are as followed:

- detail approaches used for dealing with overlapping electrical measurements
- compare results of different anomaly detection algorithms applied on electrical measurements
- (work in progress) show that multiple independent measurements improve accuracy of anomaly detection

2 Related Works

[2] is a related work that also uses electrical measurements for anomaly detection. In [2], a neuro-cognitive science inspired architecture, HTM, was developed to perform unsupervised anomaly detection. While electrical measurements (uPMU data) were used for this approach, the anomaly detection is real-time, involves hierarchical temporal memory, but does not include overlapping electrical measurements like in our case.

[4] and [1] use dynamic time warping (DTW) related to the power grid; however, [4] applies it to smart meter readings for lowering the complexity of power disaggregation. [1] applies DTW to electric utility data to cluster the electric utility net data based on the distance measure to improve operation for the next day. None of these papers apply DTW with the goal to improve anomaly detection.

However, [3] uses DTW for anomaly detection. Instead of using electrical measurements, it is used on network traffic data to detect network anomalies. The network traffic is decomposed into control and data planes. Based on the DTW distance between these two, the network activities are classified as either benign or anomalous. In summary, the related works do not address the challenge of improving anomaly detection based on overlapping electrical measurements.

3 Background

The data set used for this paper originates from Los Alamos National Laboratory's power grid from two independent overlapping electrical measurement systems: ION and HIST. The data set was collected in December 2020 and contains 744-756 ION data points and 534,686 HIST data points for each time series. HIST data is collected at a higher frequency (approximately every 5 seconds) while ION is collected at a much lower frequency (approximately every hour). A subset of measurement points are common to both systems and are therefore said to be overlapping.

To find out which measurements are overlapping, we are using the DTW algorithm. It finds the similarity between two time series by calculating the distance between them. DTW involves a non-linear optimal alignment which ensures that similar time series match each other even if they are out of phase on the x-axis or require compression or expansion. In our case DTW is used to account for the difference in frequency between the two systems as well as any potentially minor discrepancies in clock synchronization. DTW has applications in many domains such as robotics, data mining,and manufacturing [6].

The anomaly detection algorithms used in this paper to compare anomalies between the overlapping time series include autoregression, level shift, and rolling average. All of these algorithms are statistical unsupervised anomaly detection approaches. Autoregression looks at autoregressive behavior changes to detect anomalies. Level shift is often used when the data has a lot of outliers because it is not as sensitive to spikes in data. It works by taking two sliding windows and comparing their median values to detect a shift of values. Rolling average takes each value and compares it to its previous value [5].

4 Implementation and Results

4.1 Dynamic Time Warping

To determine similarity among the two different systems, the DTW algorithm is used. Due to the overhead of the DTW, the fast version was implemented. Two libraries that we considered include FastDTW and DTAIDistance's fast implementation of the DTW algorithm. After implementing both of them, we decided to continue working with DTAIDistance library due to its faster run time compared to FastDTW. Both gave similar results; however, DTAIDistance's faster run-time can be explained by the fact that it is implemented in C, which includes Cython as a dependency.

We considered three approaches to obtain a viable sample of the data:

Run Number	HIST step size	ION step size	Run-time (s)
1	100	2	2318.6
2	1000	1	616.7
3	1000	2	342.8
4	2000	2	222.6
5	3000	4	108.6
6	5000	7	65.7

Table 1: Step Size Running Times.

Run Number	Point Amount	Run-time (s)
7	100	24.3
8	200	38.6

Table 2: Point Amount Running Times.

1. step size, which includes collecting data points at different steps
2. certain amount of points, for example, collecting the first 100 points
3. range of dates, for example, a day's worth of data

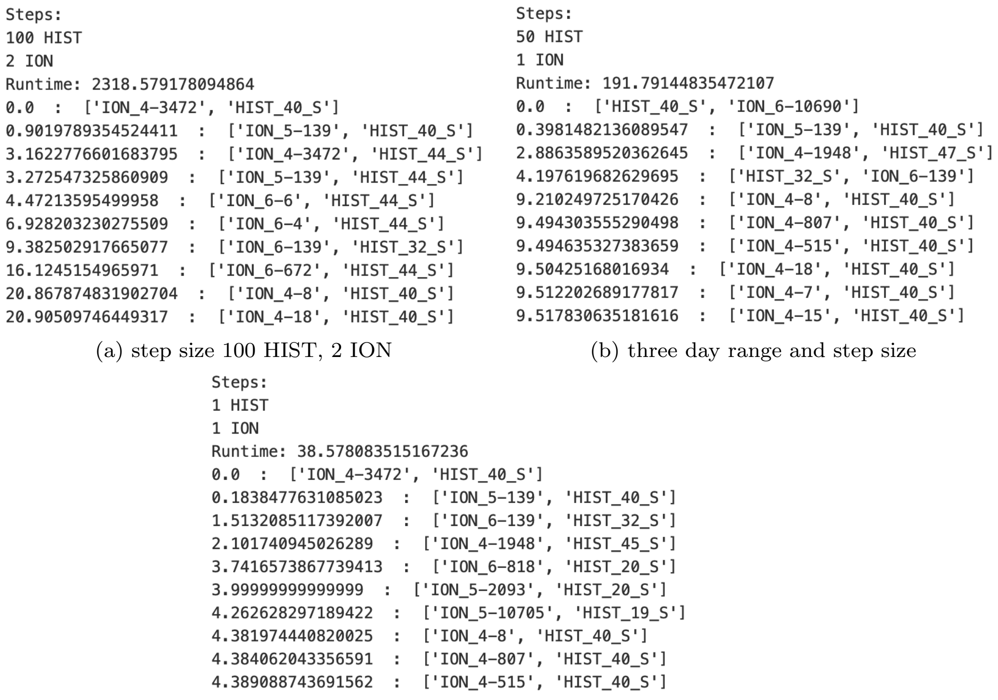
For step-size, we collected a total of 6 runs, which are depicted in Table 1. The smaller the steps, the more data points, and the longer the algorithm takes to run to compare all of the time series with the DTW algorithm. For certain amount of points, we did two runs total where we took the first 100 and 200 points as seen in Table 2. Similar to the step runs, as the amount of points increase so does the run-time.

For date range, one run was completed including data over a period of three days as can be seen in Table 3. However, this method took a long time compared to the others because the date had to be filtered beforehand which is not included in the run-time listed in the table. However, by reducing the data range we were able to reduce the step size and include more data points.

For each run, the DTW calcutes the distance for each ION and HIST pair and sorts them based on distance. The lower the distance, the more similar the data. The DTW distance results for the first 200 points stayed more stable distance-wise compared to the step size and three day range results, due to deviations in the data appearing later. The first 200 points and step size results

Run Number	Date Range	HIST step size	ION step size	Run-time (s)
9	3 days	50	1	191.8

Table 3: Date Range Running Times.



```
Steps:
100 HIST
2 ION
Runtime: 2318.579178094864
0.0  :  ['ION_4-3472', 'HIST_40_S']
0.9019789354524411  :  ['ION_5-139', 'HIST_40_S']
3.1622776601683795  :  ['ION_4-3472', 'HIST_44_S']
3.272547325860909  :  ['ION_5-139', 'HIST_44_S']
4.47213595499958  :  ['ION_6-6', 'HIST_44_S']
6.928203230275509  :  ['ION_6-4', 'HIST_44_S']
9.382502917665077  :  ['ION_6-139', 'HIST_32_S']
16.1245154965971  :  ['ION_6-672', 'HIST_44_S']
20.867874831902704  :  ['ION_4-8', 'HIST_40_S']
20.90509746449317  :  ['ION_4-18', 'HIST_40_S']
```

(a) step size 100 HIST, 2 ION

```
Steps:
50 HIST
1 ION
Runtime: 191.79144835472107
0.0  :  ['HIST_40_S', 'ION_6-10690']
0.3981482136089547  :  ['ION_5-139', 'HIST_40_S']
2.8863589520362645  :  ['ION_4-1948', 'HIST_47_S']
4.197619682629695  :  ['HIST_32_S', 'ION_6-139']
9.210249725170426  :  ['ION_4-8', 'HIST_40_S']
9.494303555290498  :  ['ION_4-807', 'HIST_40_S']
9.494635327383659  :  ['ION_4-515', 'HIST_40_S']
9.50425168016934  :  ['ION_4-18', 'HIST_40_S']
9.512202689177817  :  ['ION_4-7', 'HIST_40_S']
9.517830635181616  :  ['ION_4-15', 'HIST_40_S']
```

(b) three day range and step size

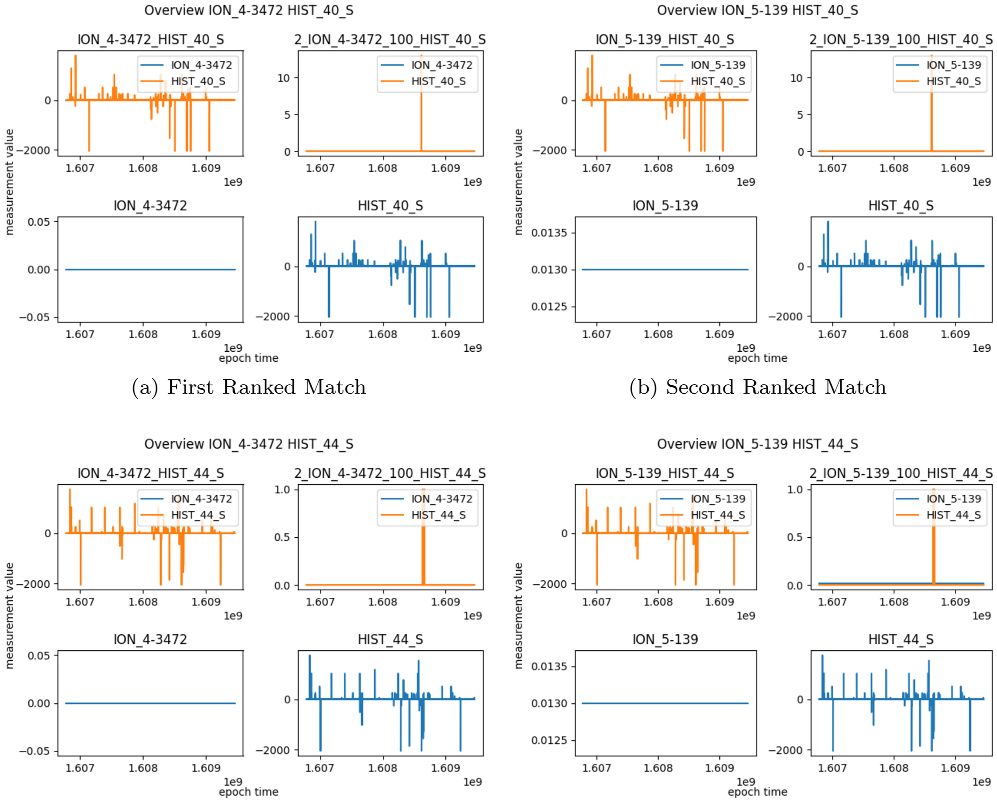
```
Steps:
1 HIST
1 ION
Runtime: 38.578083515167236
0.0  :  ['ION_4-3472', 'HIST_40_S']
0.1838477631085023  :  ['ION_5-139', 'HIST_40_S']
1.5132085117392007  :  ['ION_6-139', 'HIST_32_S']
2.101740945026289  :  ['ION_4-1948', 'HIST_45_S']
3.7416573867739413  :  ['ION_6-818', 'HIST_20_S']
3.999999999999999  :  ['ION_5-2093', 'HIST_20_S']
4.262628297189422  :  ['ION_5-10705', 'HIST_19_S']
4.381974440820025  :  ['ION_4-8', 'HIST_40_S']
4.384062043356591  :  ['ION_4-807', 'HIST_40_S']
4.389088743691562  :  ['ION_4-515', 'HIST_40_S']
```

(c) first 200 points

Figure 1: Top ten results of the different runs.

have several of the same matches and the top two results are the same. The three day results are not as similar as the other two, but still have some of the same matches in the top ten. The results of the runs are in Figure 1. At the end, we decided to settle for the step size results as an accurate similarity ranking due to its inclusion of the whole range of data and manageable run-time.

The four most similar results from the DTW 100 and 2 step size are depicted in Figure 2. Most of the points for these graphs are concentrated on y = 0, with some HIST outlier points deviating from the axis. The difference in pattern between the time series when looking at the bottom two graphs within each figure could mean one of two things, they could be different measurements; or they are the same but the ION measurements are not measured at a high enough frequency to record the spikes that appear in their HIST counterpart. The representative sample for the HIST measurements also does not include all the spikes that the actual measurements do because of the step size. This was the case with all the different methods of choosing a representative sample. The sample data is drawn in the top right corner for each graph. After we had these results, we used the four most similar matches from the DTW 100 HIST, 2 ION steps for anomaly detection.



(a) First Ranked Match

(b) Second Ranked Match

(c) Third Ranked Match

(d) Fourth Ranked Match

Figure 2: Top Four Results Step-size.
Top right corner of each graph includes the step size drawn.

4.2 Anomaly Detection

We took the top four matches of the DTW results and used autoregression, level shift, and rolling average on them. Autoregression and rolling average identified very similar anomalies. Level shift detected no anomalies, but it is less susceptible to time series containing noisy data. However, both autoregression and rolling average detected differences in anomalies among DTW matches as displayed in Figure 3 and 4. Figure 3 has anomalies mostly along its spikes. In comparison, Figure 4 is a straight line with no deviations and therefore has no anomalies. Since autoregression works by comparing its previous value to the current one, the spikes should appear one at a time which also highlights that the step size sample approach might have easily missed the spikes because they are relatively rare. This mismatch between anomalies for the same measurement could indicate certain types of data tampering. Further investigation needs to be done to understand the cause of mismatched anomalies between the HIST and ION measurements.

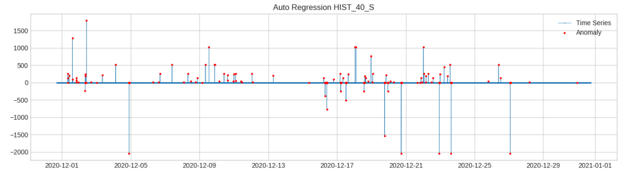


Figure 3: Autoregression HIST 40 S

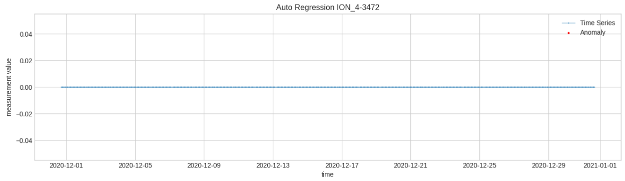


Figure 4: Autoregression ION 4 3472

5 Conclusions

The overall goal of this research is to improve the accuracy and speed of anomaly detection by using overlapping electrical measurements from the Los Alamos National Laboratory power grid and to validate the measurements. This is especially helpful when trying to detect cyber attacks because anomaly detection can be seen as an early warning system. The DTW algorithm found similar measurements that can be used for anomaly detection. The difference in frequency between the two independent system introduces challenges that are most likely impacting the DTW results and the difference in anomalies.

6 Future Work

Future work includes looking at other statistical methods that assess when the differences in anomalies are relevant enough that the measurements are not the same. In addition, exploring different ways to filter the data before running DTW to reduce run times and improve the scalability of the approach. Using the pearson correlation coefficient would be a possible option to find similar time series within each system and then filter based on the results.

Another possible area for future work includes determining which anomaly detection algorithm is the most accurate one. A possible approach would be to implement a generative model to insert and simulate an artificial anomaly and see which algorithm detects it. Or generating a synthetic data set and label it to make this a supervised anomaly detection problem which has more known methods.

References

[1] Jason R. Ausmus et al. "Improving the Accuracy of Clustering Electric Utility Net Load Data using Dynamic Time Warping". In: *2020 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*. 2020, pp. 1–5. DOI: 10.1109/TD39804.2020.9299915.

[2] Anomadarshi Barua et al. "Hierarchical Temporal Memory Based Machine Learning for Real-Time, Unsupervised Anomaly Detection in Smart Grid: WiP Abstract". In: *2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS)*. 2020, pp. 188–189. DOI: 10.1109/ICCPS48487.2020.00027.

[3] Diab M. Diab et al. "Anomaly Detection Using Dynamic Time Warping". In: *2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*. 2019, pp. 193–198. DOI: 10.1109/CSE/EUC.2019.00045.

[4] Georgia Elafoudi, Lina Stankovic, and Vladimir Stankovic. "Power disaggregation of domestic smart meter readings using dynamic time warping". In: *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. 2014, pp. 36–39. DOI: 10.1109/ISCCSP.2014.6877810.

[5] Roberto Medico. "rob-med/awesome-TS-anomaly-detection". In: (2020). DOI: 10.5281/zenodo.3972944.

[6] Stan Salvador and Philip Chan. "FastDTW: Toward accurate dynamic time warping in linear time and space". In: *KDD workshop on mining temporal and sequential data*. Citeseer. 2004.

[7] Chee-Wooi Ten, Junho Hong, and Chen-Ching Liu. "Anomaly Detection for Cybersecurity of the Substations". In: *IEEE Transactions on Smart Grid* 2.4 (2011), pp. 865–873. DOI: 10.1109/TSG.2011.2159406.

[8] Venkatesh Venkataramanan et al. "Measuring and Enhancing Microgrid Resiliency Against Cyber Threats". In: *IEEE Transactions on Industry Applications* 55.6 (2019), pp. 6303–6312. DOI: 10.1109/TIA.2019.2928495.